The relevance of metallic species for monitoring acidity in mine waters: statistical modeling for highly contaminated sites

Ana Barroso[1], Teresa Valente[1], Amélia Paula Marinho-Reis[1,2], Isabel Margarida H. R. Antunes[1]

[1] ICT – Institute of Earth Sciences, pole of University of Minho, University of Minho, Braga, Portugal
[2] GEOBIOTEC, Geosciences Department, University of Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal

Abstract

Acid mine drainage (AMD) challenges the global mining industry due to its severe environmental impact on aquatic ecosystems, water resources, and soils. Areas affected by AMD are typically characterized by high acidity, low pH (< 3), and elevated concentrations of sulfate and metals (Fe, Al, Mn, and others). Monitoring and predicting water acidity in AMD systems is critical for assessing environmental impacts and implementing effective mitigation strategies. However, the high concentrations of dissolved metals in waters affected by AMD can affect acidity levels, making them potentially suitable indicators and predictors of water acidity.

The present study introduces a statistical approach with important potential in estimating the acidity of water affected by AMD using various physicochemical parameters and metal concentrations as predictive variables. Water samples were collected in a highly contaminated system – the Trimpancho mining complex (Iberian Pyrite Belt, Spain), which comprises four abandoned medium-sized mines. Sampling was conducted during dry (February 2022) and rainy periods (October 2023). A total of 33 water samples were collected in three different hydrochemical environments: river courses, open pits, and water puddles (formed from water sources or temporally accumulated after precipitation). Field measurements included the pH and electrical conductivity, while sulfate, acidity, and metal(loid)s concentrations were determined in the laboratory. Two categorical variables were created: sampling survey, comprising two groups (wet and dry seasons), and water type, comprising three categories (stream water, water puddles, and open pit water).

Summary statistics were calculated to describe the quantitative variables. The Mann-Whitney U and Kruskal-Wallis H tests were used to identify significant differences in water properties between surveys and water types. Stepwise linear discriminant analysis (LDA) was performed to determine which variables provided the best separation between sampling environments. Multiple linear regression (MLR) was used to determine which hydrochemical variables best predicted acidity.

The samples generally exhibit extreme acidity conditions, with the pH ranging from 1.67 to 3.06. Sulfate and acidity concentrations ranged from 482 to 171 639 mg L^{-1} and 238 to 125 250 mg L^{-1} $CaCO_3$, respectively. Iron and aluminum were the dominant dissolved cations. The results showed significant differences (ρ <0.05) for some variables between dry and wet seasons, as well as between water types (ρ <0.05). The LDA model showed that Fe^{3+}, As, Zn, EC, Mn, Al, and Pb were the most important factors in discriminating the water types, with 97% of the samples correctly classified. Therefore, water puddle samples were excluded from the stepwise MLR analysis. The best MLR model suggests that Cu, Fe(total), Mn, Zn, and Al were the best predictor variables of acidity (R^2=98%) in this hydrogeochemically complex system.

In conclusion, although the dataset is relatively small, its complexity illustrates well the influence of weather conditions on the hydrogeochemistry of these waters. The results show that an MLR model including Cu, Fe, Mn, Zn, and Al provides, on average, a smaller error compared to the conventional acidity calculation method. Thus, this study highlights the importance of incorporating site specific characteristics of these complex systems to develop robust models for individual mining areas.

Keywords: Hydrochemistry; Metals concentration; Statistics; Multiple linear regression; Iberian Pyrite Belt